Supporting Information: Neural Network Potentials for Reactive Chemistry: CASPT2 Quality Potential Energy Surfaces for Bond Breaking

Quin H. Hu, Andrew M. Johannesen, Daniel S. Graham, and Jason D. Goodpaster*

Department of Chemistry, University of Minnesota. 207 Pleasant St. SE, Minneapolis MN 55455, USA.

E-mail: jgoodpas@umn.edu

1 Atomic Environment Vector Parameters

The ANI model used identical parameters and structure to the published ANI-1ccx work by Smith et al.[S1] The parameters used to calculate Atomic Environment Vector (AEV) of a system are provided below.

For radial parameters of G_m^{R},

$R_{\mathrm{C}} = 5.2$ Å,

$\eta = [16]$,

$R_{\mathrm{S}} = [0.900000, 1.168750, 1.437500, 1.706250, 1.975000, 2.243750, 2.51250, 2.781250, 3.050000, 3.318750, 3.587500, 3.856250, 4.125000, 4.39375, 4.662500, 4.931250]$ Å.

For angular parameters of G_m^{A},

R_C = 3.5 Å,

η = [8],

R_S = [0.900000, 1.550000, 2.200000, 2.850000] Å,

ζ = [32],

θ_S = [0.19634954, 0.58904862,0. 9817477, 1.3744468,

1.7671459, 2.1598449, 2.552544, 2.945243].

2 Empirical Self-energy Linear Fitting Parameters

The ANI model performs an initial atomization energy style calculation to determine an empirical self-energy term. Essentially linear fitting parameters for each atomic element is provided to determined a base energy of a given system, the actual NN would simply calculate the much smaller atomization energy, and the entire system returns the total energy as a sum. The linear fitting parameters in Hartrees used are provided below. The DFT parameters are identical to the ANI-1x work.

DFT Linear fitting parameters in Hartree:

H = -0.600952980000

C = -38.08316124000

CASPT2 Linear fitting parameters in Hartree:

H = -0.599939576600

C = -38.00037263040

3 CASPT2 Neural Networks Training Trajectory

In this work, the DFT ensembles were all trained for about 1,500-2,000 epochs on the perspective training data.

The training was terminated with RMSE performed well and as expected as shown in the data. - The RMSE plateau's like shown in the figure.

However, the CASPT2 ensembles were all trained for about 100 epochs for each NN. This decision is made to prevent potential over-fitting with small size training data. After transfer learning with more than 80 percent of parameters frozen, the DFT trained ensembles were retrained with just a couple thousands of CASPT2 data, which is significantly less than the DFT training set. The training trajectory of a NN training on CASPT2 data is provided below, where the RMSE of the NN on MEP, training data, and MD data used as validation are graphed against numbers of epochs. As shown, after the first 100 epochs, the RMSE of both MEP and validation are increasing, which is a sign of over-fitting as the NNs are trained for longer time over the small size training data.

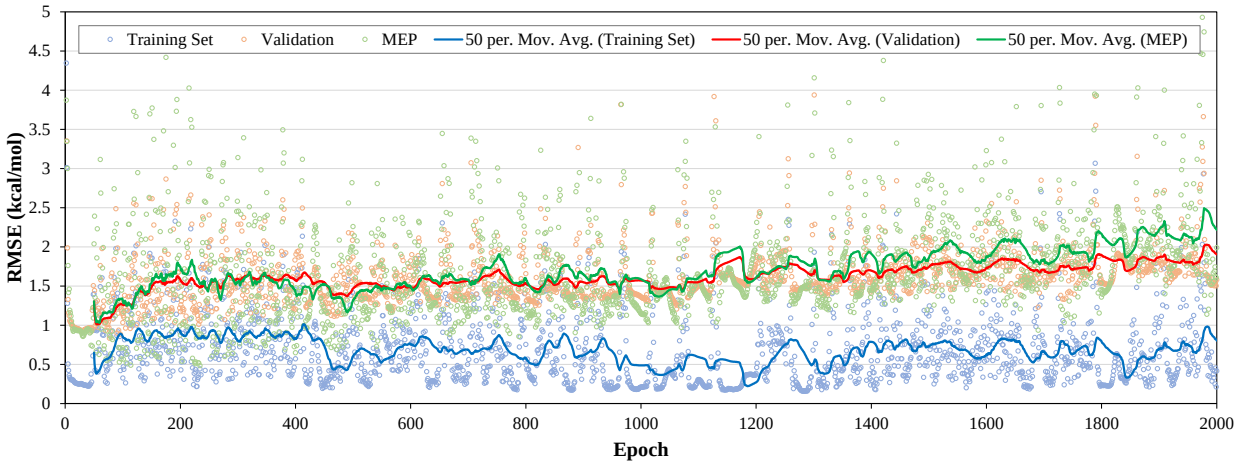

Figure S1: Training trajectory of the CASPT2 network in terms of RMSE performance with respect to Epochs. The raw RMSE are shown as dots, while the moving average per 50 points were graphed as lines.

4 CASPT2 Neural Networks Potential Transferability

Figure below showcase the comparison between the CASPT2 dissociation of C4(23) and C6(34). The exact dissociation energy are reported in Table 2 of the main text. NNP is less than 0.6 kcal/mol off from CASPT2 energy at 5 Å for C4(23), and NNP is less than 0.5 kcal/mol off for C6(34). The reference CASPT2 data shows a relative energy difference

between the dissociation energies C4(23) and C6(34) of 1.08 kcal/mol compared to 1.16 kcal/mol for the CASPT2 trained NNP. Thus, despite the C6(34) bond dissociation only being included in the DFT data, the CASPT2 trained network correctly predicts this energy due to the transfer learning.

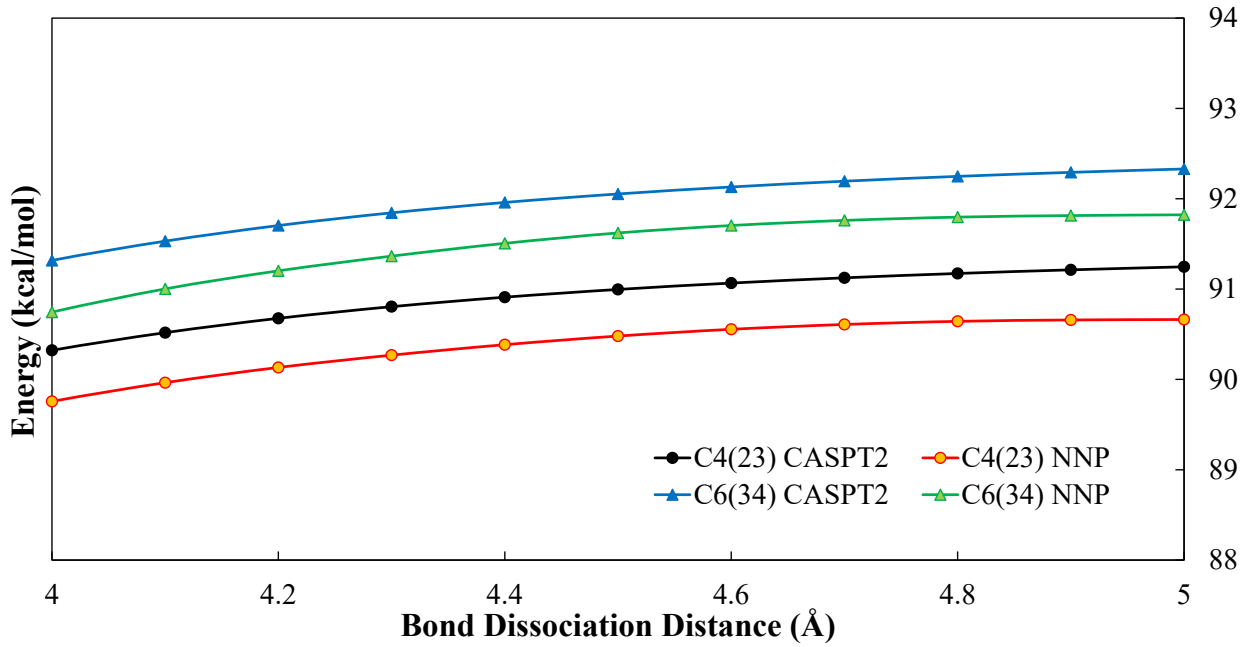

Figure S2: Potential energy comparison between CASPT2 and NNP on dissociation of C4(23) and C6(34). The round dotted lines are dissociation of butane at carbon 2 and 3; the triangle dotted lines are of hexane at carbon 3 and 4.

References

(S1) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nature Communications* **2019**, *10*, 2903.